# Training adult novices to use computers: Effects of different types of illustrations

E. James Kehoe*, Timothy C. Bednall, Leon Yin, Kirk N. Olsen, Claudia Pitts, Julie D. Henry, Phoebe E. Bailey

*School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia*

ABSTRACT

Two experiments examined the acquisition of word-processing skills (Experiment 1) and internet usage skills (Experiment 2) by novice adults using three types of illustration, specifically, full-screen illustrations with the text superimposed, icons embedded in the text, and a control, text-only condition. Training with the full-screen or embedded icons yielded significant improvements in performance relative to the text-only condition only in tests that simply required the participants to repeat the step-by-step directions. When tests required that the participants themselves retrieve and reproduce the sequence of steps, the training with illustrations – especially the full-screen illustrations – repeatedly had a negative effect on performance. The results suggest that the illustrations, by making it easy to execute each step of a procedure during training, encouraged superficial processing, which in turn yielded a set of stimulus-response associations rather than a well-linked sequence of actions in long-term memory.

© 2008 Elsevier Ltd. All rights reserved.

## 1. Introduction

Despite the prevalence of computers, there is still no shortage of adults who are largely strangers to them and find them baffling. Although graphical user interfaces are intended to appeal to familiar concepts of the "desktop," "documents," and "files," many adults find these analogies to be unhelpful and even misleading (Carroll, 1997). To assist adult novice computer users, the present experiments were aimed at testing the effectiveness of a self-paced, instructional workbook. More particularly, we assessed the effects of different types of illustrations in training. The use of adult novices was of theoretical as well as practical interest for two reasons:

First, the load on working memory during acquisition of a new skill heavily depends on a learner's existing knowledge and skills, which are stored as organised structures in long-term memory. These memory structures, variously called schemas or mental models, allow the learner to filter and fuse incoming information into single "chunks" for manipulation in working memory (e.g., Reed, 2006). As a consequence, novices, who lack the relevant schemas, are prone to overload even for elementary components of a new skill (Clarke, Ayres, & Sweller, 2005; Pollock, Chandler, & Sweller, 2002).

Second, according to the cognitive theories of instructional design (Mayer & Moreno, 2003; Paas, Renkl, & Sweller, 2003; Sweller, van Merrienboer, & Paas, 1998), efficient learning depends crucially on optimising the load on the learner's working memory; an overload causes confusion, while an underload causes boredom

and inattentiveness. Older adults may be particularly susceptible to overload, because the capacity of working memory, as well as other aspects of information processing, decline as a function of normal adult aging (Mykityshyn, Fisk, & Rogers, 2002; Van Gerven, Paas, Van Merrienboer, & Schmidt, 2000).

Despite the limitations of cognitive architecture, novices can be helped to avoid overload and jump-start schema acquisition. In general, learning will be most readily initiated by using the simplest, realistic examples that contain all the key features of the skill (Van Merriënboer, Clark, & de Croock, 2002; Van Merriënboer, Kirschner, & Kester, 2003). Moreover, the introduction to each procedure should require that only two or three elements be simultaneously held in working memory. For introducing a novice to a relatively fixed procedure (e.g., saving a computer file), the use of step-by-step instructions with minimal explanation is thought to be least likely to overload working memory and most likely to start schema formation (Carroll, 1997; Clarke et al., 2005; Pollock, Chandler, & Sweller, 2002).

For instructing novice computer users without a human trainer, the ideal method might be an interactive tutorial with on-screen pointers to icons accompanied by an auditory narrative (Lin & Hsieh, 2006; Mayer & Moreno, 2003; Van Gerven, Paas, Van Merriënboer, Hendriks, & Schmidt, 2003). As attractive as computer-based tutorials are, they pose a conundrum for self-instruction by novices: Launching and using a computer-based tutorial requires exactly the basic knowledge of computers that novices lack.

In contrast, printed workbooks are readily accessible by novices, as well as being less expensive to produce and easier to revise. Workbooks, however, have their own instructional risks. Holding information in working memory while switching attention between the

---

* Corresponding author. Tel.: +61 2 9385 2 3826.
  E-mail address: j.kehoe@unsw.edu.au (E.J. Kehoe).

workbook and the computer produces spatial and temporal separation that hinders learning ("the split-attention effect," Chandler & Sweller, 1996; Cierniak, Scheiter, & Gerjets, 2009; Ginns, 2006; Mayer & Moreno, 2003). Encoding material in the workbook and on the computer screen also rely solely on the visuospatial channel of working memory, thus further increasing the risk of overload (Mayer & Moreno, 2003).

Although printed workbooks are vulnerable to the split-attention effect, there are strategies that can increase their usefulness. Among other things, illustrations have long been used in instructional materials. Diagrams that depict physical or conceptual relationships generally enhance learning, provided that they are geared to a learner's background knowledge, minimize split-attention effects, and are not cluttered with redundant textual explanation (Carlson, Chandler, & Sweller, 2003; Vekiri, 2002, but see De Westelinck, Valcke, De Craene, & Kirschner, 2005). Similarly, a sequence of illustrations can enhance the learning and memory of assembly procedures, provided they are organised in line with the subjective structure of the task (Brunyé, Taylor, Rapp, & Spiro, 2006; Zacks & Tversky, 2003).

To test the relative value of illustrations in workbooks for novice computer users, two experiments were conducted. In each experiment, a within-subject design was used. That is, all the participants received training using a workbook in which different components entailed one of three types of illustration: (1) *full screen*, in which step-by-step, written directions were superimposed on a photo of the entire computer screen, (2) *embedded icons*, in which the relevant lines of text in the step-by-step, written directions contained illustrations of the relevant icons, and (3) *text-only*, in which the directions were presented without any illustration.

Based on the previous research and cognitive theories of instructional design, the present experiments tested the following hypotheses:

**Hypothesis 1.** Both types of illustrations should produce greater learning than the text-only condition, by helping the learner find and recognise features on the screen mentioned in the directions (Chandler & Sweller, 1996; Sweller & Chandler, 1994). According to Vekiri (2002), novice learners can be assisted by cues such as arrows, which were used in our full-screen illustrations, or embedded descriptors, as was the case with our embedded icons.

**Hypothesis 2.** Full screen illustrations could be more beneficial than the embedded icons. The full-screen illustrations with the superimposed instructions should reduce the time that a single instruction must be held in working memory by guiding the learner's eye to the relevant portion of the computer screen, thus ameliorating the split-attention effect.

**Hypothesis 3.** Alternatively, the embedded icons could be more beneficial, because they reduce the load on working memory by both directing the learner's attention to the key elements (Pollock et al., 2002), and integrating them into the text of the instruction (Mayer & Gallini, 1990; Sweller et al., 1998).

## 2. Experiment 1

Experiment 1 conducted training in word-processing using participants who had previously received a day-long introduction to the computer keyboard, mouse, and operating system. According to Clarke et al. (2005), this sequencing of general computer instruction and specific word-processing instruction minimized the cognitive load that would have occurred had participants had to acquire the two skills concurrently.

### 2.1. Method

#### 2.1.1. Participants

The final sample ($N = 41$) had a mean age of 56 ($SD = 11.2$; range: 31–80) and contained 59% females. In addition, there were 12 participants who failed to complete the experiment, largely because they found that word-processing was not to their liking. The participants all reported having little or no computer experience, could understand English to an acceptable standard, and had no apparent physical or mental disabilities. All the participants had received 6 h of instruction in basic keyboarding and the operating system several days before this experiment.

#### 2.1.2. Apparatus and materials

The experiment was conducted using computers running Microsoft Windows XP and Word. Training centered on a workbook containing two major sections. The first section contained a review of keyboarding, mouse navigation, highlighting, and scrolling. The second section contained 14 units concerning the following functions: (1) undo, (2) save files, (3) open files, (4) change view, (5) zoom, (6) cut and paste, (7) underline, (8) bold, (9) italics, (10) change the font type, (11) change the font size, (12) use spell check, (13) bullet points, and (14) print. All the units were organized around a set of click-by-click instructions in a series of recipe-like steps that rapidly engaged the participant in the task with minimal preliminary explanation (Carroll, 1997; Pollock et al., 2002). Table 1's top portion shows an example of the text for Unit 4 - Change View.

Each unit was presented in one of three formats: *text-only*, in which only the written instructions were presented; *embedded icons*, in which the key icon for a button or menu item was inserted next to its name in the text; and *full screen*, in which boxes containing the step-by-step instructions were superimposed on a photograph of the screen with arrows pointing to the appropriate icons. To obtain a measure of the cognitive load during training, the participant was asked to rate the ease-of-learning for each unit as it was completed (see Van Gog & Paas, 2008 for a discussion of different types of cognitive load measures). This rating used a six-point scale ranging from 1 ("Very Difficult – I had trouble learning this task") to 6 ("Very Easy - I had no trouble learning this task").

Participants received one of three versions of the workbook, in which the type of illustration was counterbalanced across units. Specifically, each workbook contained five units in one format

**Table 1**
Examples of text-only instructional format.

Experiment 1
*Print Layout View*
You are currently viewing the file "Microsoft Word Tutorial 1" in "Normal View.
    " However, most people prefer to view a document in "Print Layout View,
    " as this shows the document similar how it will appear in real-life on paper.
Step 1: Go to the View Menu. Select "View" from the menu at the top using the mouse.
Click on "Print Layout," the third choice on the menu. *You will notice the the document
    now looks like a piece of paper in real-life. To return to the "Normal View" setting,
    follow the same procedure selecting "Normal".*

Experiment 2
*Summary of Unit 5 - Accessing e-mail*
(1) Introduction: "Starting from this unit, we will learn how to use e-mail.
    Your teacher will have set up an e-mail account for you already.
    We will be using a web-based e-mail system called ..."
(2) Glossary: definitions of an e-mail, ID, password, and inbox.
(3) Exercise 5a - six steps for accessing Yahoo mail," e.g.,
    "Step 1 - Type www.yahoo.com.au into the address bar,
    and press Enter.
(4) Exercise 5b – three steps for checking for a new message.
(5) Exercise 5c – repeat Exercises 5a and 5b.
(6) Summary of the steps in the two sub-procedures.

(Units 1, 3, 6, 9, 12), five units in another format (Units 2, 5, 8, 11, 14), and four units in a third format (4, 7, 10, 13). Although one format was underrepresented in each workbook, the counterbalancing ensured that no single format was consistently underrepresented. Participants in any single training session were given the same workbook.

### 2.1.3. Training procedure

The training and testing were conducted in small groups, most containing four to six participants. The units were self-paced, and the training time for each participant ranged between 2 and 3 h. As the participants studied the workbook, they were allowed to ask questions. In response, the trainer initially referred to the workbook and re-phrased words from the workbook. The trainer did not point directly at the computer screen. In some cases, however, the participant needed additional assistance to avoid becoming completely frustrated.

### 2.1.4. Testing and scoring

As the participants completed the workbook, they were individually tested. It was explained to the participants that the purpose of the test was to measure the effectiveness of different teaching methods, not their individual ability. The trainer read through the instructions of the test and allowed the participants to ask any clarification questions. The participants were instructed to complete each step in sequence as quickly as they could, and to progress to the next step only after finishing the preceding step. There was no time limit for the test, which typically ran for approximately 15–30 min. The participants were not allowed to consult the workbook during the test.

The test contained 14 items. Half of the items (Set A) asked the participants to follow the step-by-step written directions for a procedure. For example, the test item for Unit 2 read: "Step 1: go to the File menu; Step 2: select 'Save'; Step 3: enter a file name; Step 4: click Save." The other half of the items (Set B) asked the participants to generate the steps needed to accomplish a procedure. For example, the test item for Unit 7 read: "Underline the following text".

During the test, the participant's actions on the computer screen were recorded using HyperCam (Version 2.11). An action was recorded whenever the participant either clicked the mouse button, pressed a key, or moved the mouse over a menu item for at least one second. Based on these records and those of the trainer, two key dependent variables were recorded. First, the time taken to complete each test item was defined as the interval from the end of the last action on the previous item to the end of the assigned item. Second, errors were defined to include (1) irrelevant actions, (2) actions out of sequence, (3) failures to complete an item, and (4) verbal requests for help. Following testing, all participants were asked to rate each type of illustration on a five-point rating scale ranging from "Strongly Disliked" to "Strongly Liked."

### 2.1.5. Statistical analyses

Statistical analyses were conducted using three contrasts in which the Text, Icon, and Screen conditions were compared in a pairwise fashion (O'Brien & Kaiser, 1985). Because these comparisons were not orthogonal, a Bonferroni adjustment was applied to guard against inflation of Type I errors. Specifically, the conventional Type I error (.05) was reduced to .05/3 = .017. Although the $F$ ratio is robust in the face of non-normality and heterogeneity of variance, this reduction also helped guard against any inflation of the Type I error from those sources (Lindquist, 1953).

As a measure of effect size, partial eta squared ($\eta_p^2$) was adopted (Cohen, 1973). This measure equals the proportion of explained variance: $SS_{effect}/(SS_{effect} + SS_{error})$. According to Cohen (1988), $\eta_p^2 = .010$, .059, and .138 imply small, medium, and large

effects, respectively. Finally, means are accompanied by a figure in parentheses that represents the standard error-of-the-mean (± SEM).

### 2.2. Results

Preliminary analyses of the dependent variables were conducted to determine whether there were any discernible age-related or gender-related correlations. None of these analyses yielded any significant results (all $p > .10$). Furthermore, the preliminary analyses revealed that errors and completion times were positively correlated ($r = 0.51$). Generally, however, there were few errors, averaging 1.0 (± .3) per test item. In contrast, the participants took considerable time to complete their tasks, averaging 85 s (± 11) per test item.

Item-by-item analyses revealed similar patterns in the errors and completion times. However, the effect sizes for the errors were far smaller than the corresponding effect sizes for completion times. Among the 14 items, "large" effects ($\eta_p^2 > .138$) were obtained in the pattern of errors for only 4 items, but large effects in completion times emerged for 10 items. Consequently, presentation of the detailed results will focus on completion times.

#### 2.2.1. Completion times

Fig. 1 summarizes the mean completion times for the items in Set A (left panel) and Set B (right panel). As can be seen, there was a large interaction between the type of illustration and the testing sets. When the participant was asked simply to follow the steps of a written procedure (Set A), prior training with the text-only instructions yielded the longest completion times, while training with the embedded icons and full-screen illustrations led to progressively shorter completion times. In contrast, when the participants were required to retrieve the steps of a procedure (Set B), the reverse pattern appeared. That is, the text-only training produced the shortest completion times, and the full-screen training led to the longest times.

The statistical tests confirmed this pattern. While the absolute completion times are depicted, the statistical tests were based on the average of the standardized scores for the contributing items to ensure that they received equal weightings. The pairwise comparisons (Text vs. Icon, Text vs. Screen, Icon vs. Screen) significantly interacted with the contrast for testing set (A vs. B), $F$ (1, 40) = 43.71, 15.30, 15.27, $MSE$ = .272, all $p < .01$, $\eta_p^2$ = .52, .27, .26, respectively.
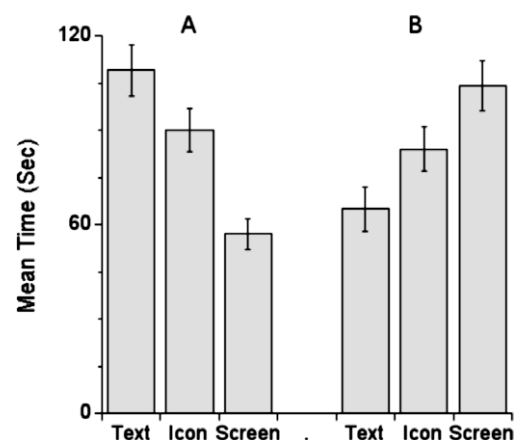


**Fig. 1.** Mean completion times in Experiment 1 as a function of the type of illustration (Text, Icon, Screen) and test task (A and B).The error bars indicate one standard error-of-the-mean.

In order to identify the possible source of this interaction, the completion times for the individual test items were examined. Fig. 2 shows the mean completion times for the items in Set A (left panel) and Set B (right panel). With some exceptions, the pattern seen in the overall analysis appeared in the results for the individual items. For the items in Set A, training with the full-screen illustration consistently produced relatively short mean completion times ($M = 56 \pm 5$ s). The text-only and embedded icon training produced a greater range of results, which nevertheless tended to be longer than those produced by the full-screen training. Across Set A's items, the means for the text-only and embedded icon conditions were 113 s ($\pm 8$) and 97 s ($\pm 7$), respectively. Within the Set B items, the text-only training tended to be shorter ($M = 66 \pm 7$ s) than those produced by training with embedded icons ($M = 84 \pm 7$ s) and full-screen illustrations ($M = 97 \pm 8$ s).

Among the individual items, the exceptions to the general trends are easy to see. Several items in both sets were relatively insensitive to the training conditions. These items generally produced relatively short completion times. Among the items that were sensitive to the previous training, three stood out, because the embedded icon training led to longer completion times than either the text-only or full-screen training. These items were Cut in Set A, Bullets in Set B, and Undo in Set B.

To test the reliability of these item-by-item differences, post-hoc comparisons were conducted using a Scheffé criterion, $F_{critical}$ (1, 38) = 9.65, $p < .004$). These tests confirmed that, for Cut and Undo, training with the embedded icons led to significantly longer times than either the text-only or full-screen training. The apparent large differences for Bullet were accompanied by considerable variability. Thus, the mean for the embedded icon training failed to differ significantly from either the full-screen training, $F(1, 38) = 7.19$, $MSE = 5995$, $p = .011$, $\eta_p^2 = .16$, or text-only training, $F(1, 38) = 5.69$, $MSE = 5995$, $p = .022$, $\eta_p^2 = .13$.

Among the other items, the post-hoc tests confirmed that, in Set A, the full-screen training yielded significantly shorter completion times than the text-only training for Fontsize and Bold. In Set B, the full-screen training yielded significantly longer completion times than the text-only training for Italic and Spell-Check. The tests for the embedded icon training yielded intermediate outcomes, sometimes differing from the text-only training, sometimes from the full-screen training, and sometimes from both.

### 2.2.2. Ratings for ease-of-learning and liking

To the extent that the ease-of-learning ratings are inversely related to the cognitive load experienced by the participants during the training (Van Gog & Paas, 2008), the learning materials, as intended, imposed a low load. That is, the participants rated all the procedures to be easy to learn; 79% of all ratings fell in the "easy" categories (4–6). These ease-of-learning ratings had tiny negative correlations with errors ($r = -0.04$) and completion times ($r = -0.16$). Across illustration conditions, the ratings hovered around a mean of 4.4 ($\pm 0.2$). In both the overall and item-by-item analyses, any apparent differences did not even approach a liberal criterion for statistical significance (all $p > .10$). Nevertheless, the participants did express some preference for the illustrations. Specifically, ratings of liking for the full-screen illustrations ($M = 4.2 \pm 0.1$) and embedded icons ($M = 4.3 \pm 0.1$) were both significantly greater than that of the text-only condition ($M = 3.5 \pm 0.1$), $F$ (1, 40) =10.27, 18.18, $MSE = .865$ all $p < 0.01$, $\eta_p^2 = 0.20$, 0.31, respectively.

### 2.3. Discussion

The results of Experiment 1 did not support any of the hypotheses that training with the illustrations would always facilitate learning. Instead, the relative benefits from illustrations depended on the procedure being learned and the use to which the learning was put. That is, training with illustrations, especially the full-screen illustrations, tended to reduce completion times when the application of the learning entailed simply following the sequence of instructions following the training (Set A). This positive effect of the illustrations may seem inconsequential, but recall the overload one can experience in attempting to decipher and follow a set of instructions found in a manual or on a "help" screen, i.e., the immediate consequences of split-attention on performance. From this perspective, the illustrations produced a type of learning that reduced the negative effects of split-attention on performance.

In contrast, when the participants were asked to apply their learning to achieve a specified goal (Set B, e.g., "Print ..."), the pattern was reversed. The text-only training tended to produce the shorter completion times. Hence, the illustrations yielded an encoding in memory that was not conducive to retrieval of the sequence of actions for a procedure.

In the present study, the basic workbook was designed to minimize the likelihood of overloading working memory so that a novice could learn with little or no assistance. The low number of errors during the tests and the high ratings for ease-of-learning suggest that that aim was generally achieved. Hence, to the extent that the text-only version of the workbook enabled effective learning, the illustrations may have been largely unnecessary. Along this line of thought, previous research has indicated that illustrations are beneficial when the learning task imposes a high load on working memory, but they diminish in their effect when the learner has more experience in a domain and presumably a lower load on
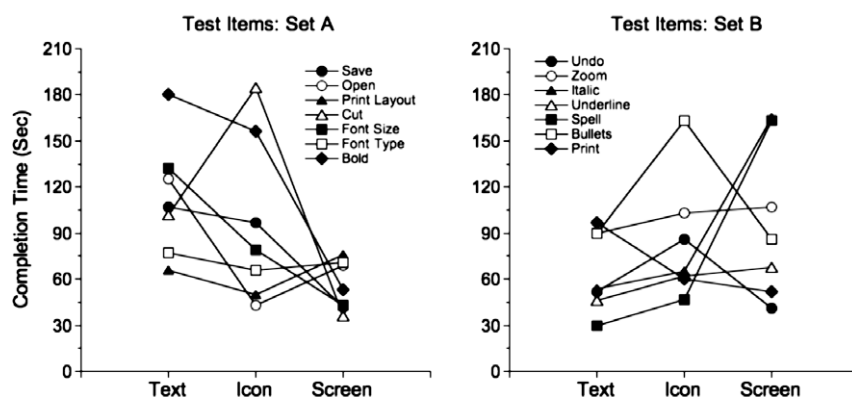


**Fig. 2.** Mean completion times in Experiment 1 as a function of the type of illustration (Text, Icon, Screen) and for each item in test task (A and B). The error bars indicate one standard error-of-the-mean.

working memory (Chandler & Sweller, 1996; Sweller & Chandler, 1994). In the present case, the load on working memory was lowered by the design of the workbook rather than by relying on the accumulation of experience.

When working memory is not overloaded, the value of instructional aids can change. The positive value of worked examples relative to problem-solving can reverse itself as even modest expertise in a domain is acquired (an "expertise reversal effect," (Kalyuga, Ayres, Chandler, & Sweller, 2003; Kalyuga, Chandler, Tuovinen, & Sweller, 2001; Reisslein, Atkinson, Seeling, & Reisslein, 2006). Similarly, explanations that are too easy can lead to suboptimal learning by loading working memory with redundant information and/or causing superficial processing of new information (Kalyuga et al., 2003; Sweller, 2005; Wittwer & Renkl, 2008).

In the present case, the illustrations may have produced superficial processing, in which each step of a procedure tended to be learned as a separate stimulus-response association rather than as a linked sequence of actions. This stimulus-response type of learning explains how training with the illustrations reduced the time needed to complete the successive steps specified in the Set A tests. Conversely, training with the text-only condition was more likely to produce storage of an integrated procedure. If the storage was integrated, the presentation of the direction for each step in the Set A tests may have introduced redundant cues that disrupted that the smooth retrieval of the integrated actions. Conversely, for the Set B items, which only provided a single, general cue, the retrieval of the integrated actions acquired in the text-only condition could proceed without disruption, while the poorly-linked stimulus-response associations acquired in training with illustrations would take longer to retrieve.

## 3. Experiment 2

Although it would seem unlikely that an expertise reversal effect occurred in Experiment 1, the participants nevertheless had had several hours of computer training prior to starting the experiment. Hence, the ease with which they learned the word-processing tasks may have partly reflected that previous learning and diluted the effects of the added illustrations. Accordingly, Experiment 2 was conducted to retest the original hypotheses using novices who received only a brief (20 min) introduction to basic keyboard and mouse skills prior to the experiment (cf. Clarke et al., 2005).

In addition to the reduction in preliminary training, a larger, more homogenous sample was recruited. Specifically, to attract more participants, the content of the training was changed from word-processing to internet usage. Discussions with participants in Experiment 1 suggested that there would be more interest in learning about internet usage. In fact, recruiting and retaining participants for Experiment 2 proved far easier than in Experiment 1. Moreover, although Experiment 1 did not reveal any age-related effects, the age range of the participants was restricted to older individuals (>54 years) as the subpopulation of adults most likely to be novices to computers and susceptible to cognitive overload.

Finally, three changes were made in the procedure aimed at reducing the variability in the training and expanding the testing:

First, the time on each unit of training was equated more tightly than in Experiment 1. Although Experiment 1 was entirely self-paced, there was an urgency felt by some slower learners to keep up with seemingly faster learners, which proved frustrating for the slower learners. Conversely, other participants proceeded so slowly that the time allocated to the computer room expired. In order to regulate the time devoted to learning, the present experiment was divided into 20 min study blocks, one for each of eight units.

Second, to give all participants a chance to review the workbooks and to give slower learners a chance to complete any material that they missed during the earlier study blocks, there was a 25 min study period just before the test. This period was also used to determine whether the explicit engagement of metacognitive processes might discourage superficial processing and facilitate the effects of illustrations (e.g., Cuevas, Fiore, & Oser, 2002). Hence, as a metacognitive activity, half the participants who received each workbook were asked to list the topics that needed the most work and then focus on those topics. The other participants were given a general instruction to review the material as they wished.

Third, both sets of tests in Experiment 1 asked the participants to reproduce exactly the procedure described in the workbook. Similarly, in Experiment 2, items like those of Set B were included to test exact reproduction of an entire procedure from memory. For example, when the participants had been trained to search for website X in the workbook, the participants were tested by asking them to search again for website X. In addition, the participants tested whether they could make simple generalizations. For example, when the participants had been trained to search for a particular website in the workbook, the corresponding generalization item asked the participants to search for a different website.

### 3.1. Method

#### 3.1.1. Participants

The participants were recruited by newspaper advertisement. The final sample ($N = 78$) had a mean age of 68 ($SD = 7.1$; range: 55–83) and contained 66% females. In addition, there were 4 individuals who did not complete training and testing.

#### 3.1.2. Materials

The workbook contained a two-page introduction to basic features and terms, i.e., monitor, keyboard, mouse, pointer, click, double click, right click, desktop, and windows, plus two preliminary units on going to a web address and browser navigation. Then, there were six main units covering the following procedures: (1) searching the internet, (2) book-marking favorites, (3) accessing e-mail, (4) composing an e-mail, (6) formatting an-e-mail message, and (6) accessing the address book.

Each unit contained a description of the task, a glossary of key terms, three exercises that provided a step-by-step description of the procedures, and a final list of the steps. As an example, the lower portion of Table 1 summarizes the unit concerning access to e-mail.

There were three versions of the workbook. The two preliminary units were always presented in a text-only format, and their test results were not used in the data analysis. Subsequently, the types of illustration were counterbalanced across pairs of the main units, specifically,1–6.Thus, all participants were exposed to all three types of illustration, specifically, text-only, embedded icons, and full-screen illustrations.

#### 3.1.3. Training and testing procedure

For each set of participants, training and testing sessions were held on a single day. The number of participants in each session ranged from four to ten participants.

Participants were free to study each unit at their own pace within a 20 min period. In a few cases, extra minutes were allowed. The participants who wanted even more time were told that they could use the study period at the end of training.

At the start of the study period, metacognitive activities were encouraged in half the participants by giving them a sheet asking them to write a "study plan" by listing topics they had found most difficult. As a control, the other half of the participants received a

sheet for making further notes plus general instructions to review the material.

The test consisted of 14 items, which, like Set B in Experiment 1, asked the participant to retrieve the actions needed to accomplish a task. For example, one item consisted of the following: "Load Internet Explorer, and go to the www.dailytelegraph.com.au website." Table 2 lists the text of the items.

In the test, Items 1 and 2 were treated as practice items (PR). Items 3–8 were labeled as Set EX, standing for "exact," because their content corresponded exactly to items in the workbook. Items 9–14 were labeled as Set GEN, standing for "stimulus generalization," because their superficial content differed from the corresponding items in the workbook while the underlying procedure remained unchanged. There was one item in each set corresponding to one of the six main units in the workbook. The test items were each presented on the screen for the participant to read. The participants were also given a sheet listing all the test items to consult.

The webpages in the test were programmed so that only correct actions produced a result. Thus, no time during testing was spent on error recovery (Carroll, 1997). To avoid the participants becoming stalled on any step, they could receive a hint by pressing the F1 key after 15 s had elapsed since completion of the previous step. When the F1 key was pressed, a box appeared directing them to the next step in the procedure. The hint box also appeared automatically if the participant waited for more than 60 s. For each test item, the computer recorded four variables: (1) time to complete the item, (2) the number of clicks on incorrect locations, (3) the number of participant-generated requests for a hint, and (4) the number of computer-generated hints. Each of the latter three variables was counted as an "error."

### 3.2. Results

Preliminary analyses revealed that completion times and errors were again positively correlated ($r$ = .45). Completion times were similar to those seen in Experiment 1 ($M$ = 89 ± 7 s), but there were more errors ($M$ = 3.5 ± 7). Analyses based on the average of the standardized scores across items failed to reveal any significant effects of illustrations for completion times or errors. There were moderate overall positive correlations of age with completion times ($r$ = .38) and errors ($r$ = .39), which the within-subject design balanced across illustration conditions in training.

Errors showed some effects unrelated to the type of illustration. Specifically, the EX tasks ($M$ = 2.4 ± 0.4) produced significantly fewer errors than the GEN tasks ($M$ = 4.5 ± 0.5), $F$ (1, 77) = 46.94, $MSE$ = 11.01, $p$ < .01, $\eta_p^2$ = 0.38. In turn, this effect interacted with the instructions at the start of the study period aimed at manipulating metacognitive activities, $F$(1, 77) = 4.24, $MSE$ = 18.52, $p$ < .05, $\eta_p^2$ = 0.05. For the EX tasks, participants who received instructions to formulate a study plan ($M$ = 2.4 ± 0.5) showed virtually the same number of errors as the control participants given only general instructions ($M$ = 2.5 ± 0.6) For the GEN tasks, however, the instructions to formulate a study plan ($M$ = 3.8 ± 0.5) yielded fewer errors than the general instructions ($M$ = 5.1 ± 0.9). The study instructions, however, did not significantly interact with the illustration conditions (all $p$ > 0.10).

Item-by-item analyses, like those conducted in Experiment 1, revealed some significant, medium-sized effects of illustrations in tests for three of the six procedures taught in workbook. The results for these tests are shown in Fig. 3.The left-hand panels show the mean completion times, and the right-hand panels show the mean number of errors for the EX and GEN test item for each procedure.

Inspection of Fig. 3 reveals that, for book-marking a website, participants trained with the embedded icons required signifi-

cantly more time to complete both the EX and GEN items than either the text-only or full-screen screen illustrations, $Fs$ (1, 72) ⩾ 3.89, $MSE$ = 11180 $p$ < .05, $\eta_p^2$ = 0.05. With regard to errors, the EX item uniformly yielded about two errors, while the GEN item yielded a progressively larger number of errors for the text-only, embedded icon, and full-screen illustrations. Statistical comparisons yielded significant interactions indicating that, for the GEN item, the text-only condition yielded fewer errors than the full-screen illustration, $F$ (1, 72) = 4.50, $MSE$ = 25.74 $p$ < .05, $\eta_p^2$ = 0.06.

In accessing e-mail, completion times for the EX item were similar across illustration conditions, taking a significantly shorter time relative to the GEN item, $F$ (1, 72) = 76.21, $MSE$ = 21744, $p$ < .01, $\eta_p^2$ = 0.51.For the GEN item, the text-only and embedded icon training produced times that were a third less than those of the full-screen illustrations. The GEN test of text-only versus full-screen illustrations attained significance, $F$ (1, 72) = 4.35, $MSE$ = 21744, $p$ < .05, $\eta_p^2$ = 0.06, while the mean completion time for GEN test after embedded icon training was intermediate and did not significantly differ from either of the other conditions. With regard to errors, there were generally fewer than four, and any apparent effects of illustration and/or type of test were not significant (all $p$ > .05).

For accessing the address book, training with embedded icons tended to produce shorter completion times and fewer errors than the other two types of illustration. The statistical pattern, however, varied. Across the EX and GEN items, the embedded icons produced significantly shorter times than the text-only condition, $F$ (1, 72) = 4.34, $MSE$ = 24330 $p$ < .05, $\eta_p^2$ = 0.06. The full-screen illustrations produced intermediate times that failed to differ significantly from either of the other types of illustration. In the case of errors, the embedded icons produced significantly fewer errors

**Table 2**
Text of test items in experiment 2

| Item | Set | Text |
|---|---|---|
| 1 | PR | Load Internet Explorer, and go to the www.dailytelegraph.com.au website. |
| 2 | PR | Go to the staff profile for Peter F. Smith and then return to the list of staff. |
| 3 | EX | Conduct an Internet search, using the key words "Desperate Housewives". Go to the first "Desperate Housewives" site that you find. |
| 4 | EX | You are on the Daily Telegraph webpage. Add it to your list of Favorites. |
| 5 | EX | You are in Yahoo Mail. Go to your Inbox, and read the 'Yahoo-Welcome' message. |
| 6 | EX | Write an e-mail message to a friend. In your message, write 'Hello'. You may send the message to any e-mail address. |
| 7 | EX | You are writing an e-mail message, and you want to change the appearance of the text. Make the word 'bold' bold, and italicise the word 'italics'. |
| 8 | EX | Add Tim Bednall (e-mail: tcbednall@yahoo.com.au) to your list of contacts. |
| 9 | GEN | Search the Internet for the movie "Superman Returns". Go to the first website you find. |
| 10 | GEN | Add the "Superman Returns" site to your Favorites. Return to the Daily. Telegraph website by clicking on it from your list of Favorites. |
| 11 | GEN | Log into Yahoo Mail, go to your Inbox, and read the e-mail from PADI. (Yahoo ID: tbednall, Password: qwerty. |
| 12 | GEN | You are looking at your Inbox. Reply to Editor Bob's message. In your message, write 'Hello' and send it. |
| 13 | GEN | You are writing an e-mail message. Format the text, so that the word 'supercalafragilisticespialladocious' is in bold, and coloured. |
| 14 | GEN | You have written a message. Send it to your friend Amanda, and CC it to your acquaintance Ekim. Retrieve their e-mails from your list of contacts. |

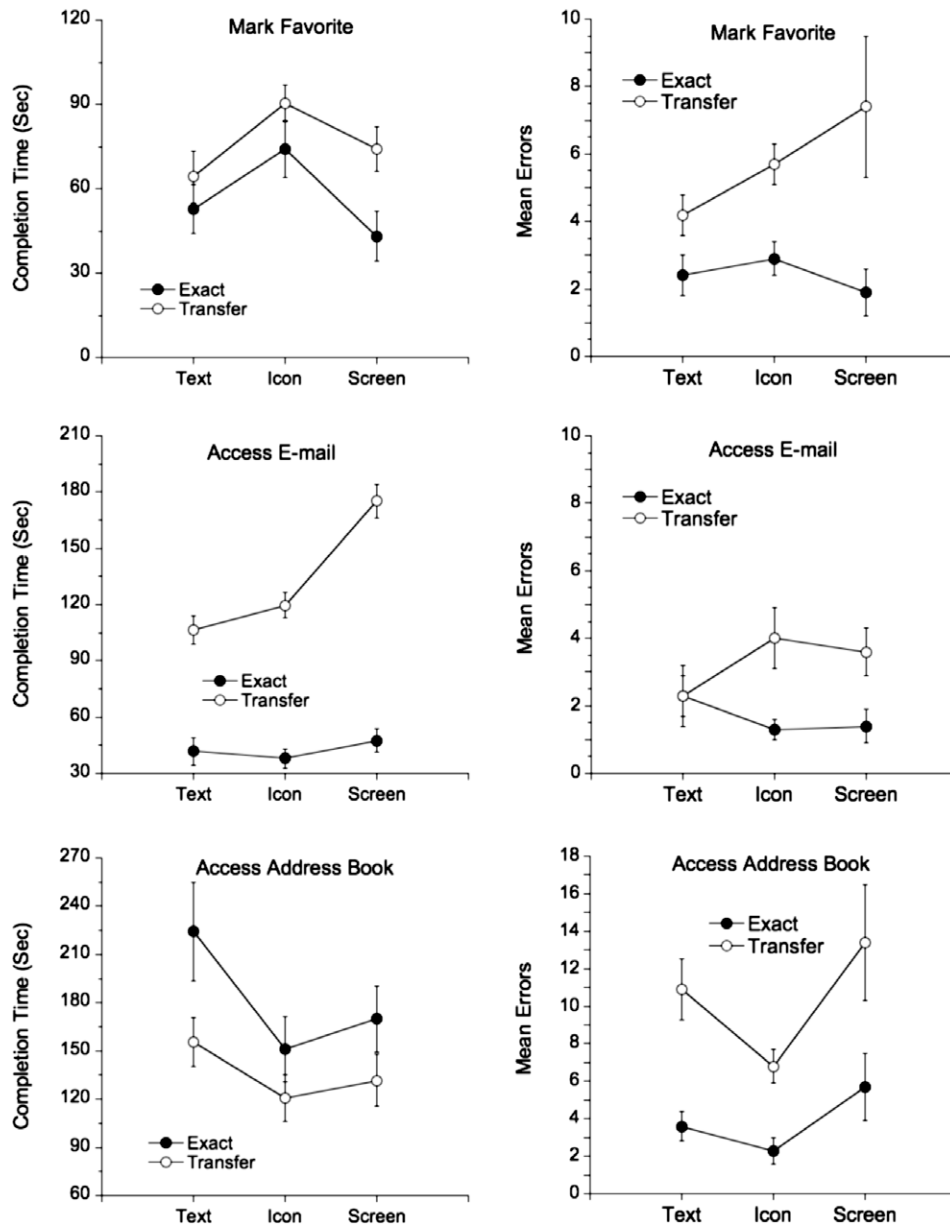PR, practice items, EX, Exact repetition from instruction, GEN, generalization items.

**Fig. 3.** Mean completion times and errors in Experiment 2 as a function of the type of illustration and test task (EX, GEN). The error bars indicate one standard error-of-the-mean. There are separate panels for the tests of the procedures for marking favorites (top row), accessing e-mail (middle row), and accessing the address book (bottom row).

than the full-screen illustrations, $F(1, 72) = 5.32$, $MSE = 120.22$ $p < .05$, $\eta_p^2 = 0.07$, while the text-only condition produced an intermediate, nonsignificant difference.

Among the other three procedures (not pictured), the tests for searching the internet, required about a minute ($M = 69 \pm 6$ s) and yielded about two errors ($M = 2.1 \pm . 4$). For both composing and formatting e-mail messages, there was a mean completion time of $82 \pm 10$ s and a mean of $2.5 \pm .6$ errors. Any apparent differences among the types of illustration were small and unsystematic (all $p > .10$).

The subjective ratings for the types of illustration were similar. The rating of liking hovered around a mean of 4.4 (all $p > .10$). As in Experiment 1, the participants generally rated the instructions to be easy; 75% of all ratings fell in the "easy" categories. The participants were also asked whether they would recommend the training course to a friend. All but one participant (98%) said they would.

### 3.3. Discussion

Despite attempts to minimize the previous computer experience of the participants, reduce the variability in training, and expand the range of testing, the results of Experiment 2 failed to support the original hypotheses concerning the beneficial effects of illustrations. Instead, the results were broadly consistent with those of Experiment 1, specifically its Set B retrieval results. That is, the participants generally rated the material to be easy to learn and, with some exceptions, made relatively few errors. Training with illustrations tended to yield longer completion times and more errors than the text-only training. Among the statistically significant effects, training with the full-screen illustration led to relatively poor performances in all three reported procedures. Specifically, training with the full-screen illustration produced the most errors in the GEN item for book-marking, the longest completion times in the GEN item for accessing e-mail, and the most

errors in the GEN item for accessing the address book. Training with embedded icons had a mixed effect. It led to the longest completion times in both the EX and GEN items for book-marking, but the fewest errors in the GEN item for accessing the address book.

The explicit engagement of metacognitive activities during the study period prior to testing did appear to lead to deeper processing of the training material. This metacognitive uniformly improved performance on the generalization tests, regardless of illustration condition.

## 4. General discussion

The present experiments were aimed at testing the added value of illustrations in learning computer applications by adult novices. In the present experiments, both positive and negative effects of illustrations were seen, depending on what was asked in the test. On the one hand, the illustrations aided the learning needed to follow the sequence of textual instructions (Set A, Experiment 1). On the other hand, when the tests in both Experiment 1 (Set B) and Experiment 2 (EX, GEN) required retrieval of the whole procedure, training with illustrations rarely improved test performance and, in repeated instances, had a negative effect.

The present results compliment redundancy effects, in which additional information interferes with learning. In the prototypic case, interference occurs when a self-explanatory diagram is accompanied by text re-describing the diagram (Chandler & Sweller, 1991) or an explanatory text is accompanied by a self-drawn diagram (Leutner, Leopold, & Sumflet, 2009) According to cognitive load theory, providing the same information in two different forms will impose an extraneous load on working memory, even if only to ascertain that the two forms are redundant (Sweller, 2005).

The present results resembled this diagram/text redundancy insofar as the addition of illustrations to self-contained text tended to hinder learning. However, unlike redundancy effects, the illustrations provided new information about the appearance of the key icons and, in full-screen illustrations, information about the spatial location of the key icons. Moreover, the addition of the illustrations did not appear to overload working memory. The participants either preferred the illustrations to the text-only condition (Experiment 1) or showed similar ratings for the three formats (Experiment 2). More generally, as intended in the use of step-by-step instructions with minimal explanation (Carroll, 1997; Pollock et al., 2002), the participants in both experiments found the procedures easy to learn. Moreover, the number of errors in most test items in both experiments was relatively low, indicating the procedures had been learned, even if their retrieval tended to be slow.

Instead of constituting extraneous information, the illustrations may have produced a load in working memory that was germane to learning, but not an optimal type of learning. By facilitating the participants' search for the key icons on the screen, the illustrations may have fostered a superficial processing of the instructions. Consequently, the participants may have more effortlessly executed the step-by-step instructions and consequently failed to store them in a strongly connected form. Rather, they were stored as weakly connected set of stimulus-response associations. Subsequently, when recall of the entire sequence was required, the participants were less able to quickly retrieve the successive steps.

In conclusion, the workbook generally led to effective learning in adult novices by minimizing the load on working memory through the use of step-by-step text instructions with minimal explanation. A diagram/text redundancy effect was avoided by using illustrations that added information not available in the text. However, the type of processing in working memory produced by illustrations made a difference to what was learned. The ease of

following the illustrated instructions may have produced superficial processing that yielded a poorly connected set of stimulus-response associations that were useful in repeating the separate steps of the procedure but not in their speedy retrieval as an integrated sequence. For findings concerning the relative effectiveness of video and static illustrations in instruction, see Arguel and Jamet (2009).

## References

Arguel, A., & Jamet, E. (2009). Using video and static pictures to improve learning of procedural contents. *Computers in Human Behavior, 25*, 354–359.

Brunyé, T. T., Taylor, H. A., Rapp, D. N., & Spiro, A. (2006). Learning procedures: The role of working memory in multimedia learning experiences. *Applied Cognitive Psychology, 20*, 917–940.

Carlson, R., Chandler, P., & Sweller, J. (2003). Learning and understanding science instructional material. *Journal of Educational Psychology, 95*, 629–640.

Carroll, J. M. (1997). Toward minimalist training: Supporting the sense-making activities of computer users. In M. A. Quiñones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace. Applications of psychological research* (pp. 303–328). Washington, DC: American Psychological Association.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332.

Chandler, P., & Sweller, J. (1996). Cognitive load while learning to use a computer program. *Applied Cognitive Psychology, 10*, 151–170.

Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior, 25*, 315–324.

Clarke, T., Ayres, P., & Sweller, J. (2005). The impact of sequencing and prior knowledge on learning mathematics through spreadsheet applications. *Educational Technology Research and Development, 53*, 15–24.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement, 33*, 107–112.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cuevas, H. M., Fiore, S. M., & Oser, R. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science, 30*, 433–464.

De Westelinck, K., Valcke, M., De Craene, B., & Kirschner, P. (2005). Multimedia learning in social sciences: Limitations of external graphical representations. *Computers in Human Behavior, 21*, 555–573.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16*, 511–525.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem-solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579–588.

Leutner, D., Leopold, C., & Sumflet, E. (2009). Cognitive load and science text comprehension: Effects of drawing and imagining text content. *Computers in Human Behavior, 25*, 284–289.

Lin, D.-Y. M., & Hsieh, C.-T. J. (2006). The role of multimedia in training the elderly to acquire operational skills of a digital camera. *Gerontechnology, 5*, 68–77.

Lindquist, E. F. (1953). *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*, 715–726.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*, 43–52.

Mykityshyn, A. L., Fisk, A. D., & Rogers, W. A. (2002). Learning to use a home medical device. Mediating age-related differences with training. *Human Factors, 44*, 354–364.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin, 97*, 316–333.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1–4.

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction, 12*, 61–86.

Reed, S. K. (2006). Cognitive architectures for multimedia learning. *Educational Psychologist, 41*, 87–98.

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16*, 92–103.

Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (pp. 147–158). NewYork: Cambridge University Press.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*, 185–233.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Van Gerven, P. W., Paas, F., Van Merriënboer, J. J., Hendriks, M., & Schmidt, H. (2003). The efficiency of multimedia learning into old age. *British Journal of Educational Psychology, 73*, 489–505.

Van Gerven, P. W. M., Paas, F. G. W. C., Van Merrienboer, J. J. G., & Schmidt, H. G. (2000). Cognitive load theory and the acquisition of complex cognitive skills in the elderly: Towards an integrative framework. *Educational Gerontology, 26*, 503–521.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist, 43*, 16–26.

Van Merriënboer, J. J. G., Clark, R., & de Croock, M. B. (2002). Blueprints fo complex learning: The 4C/ID-Model. *Educational Technology Research and Development, 50*, 39–64.

Van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38*, 5–13.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*, 261–312.

Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43*, 49–64.

Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology: Applied, 92*, 88–100.